



WORKING PAPER SERIES

Testing Asset Pricing Models with Euler Equations: It's Worse Than You Think

Christopher Neely

Working Paper 1995-018A
<http://research.stlouisfed.org/wp/1995/95-018.pdf>

Original version, 94-010A

FEDERAL RESERVE BANK OF ST. LOUIS

Research Division
411 Locust Street
St. Louis, MO 63102

The views expressed are those of the individual authors and do not necessarily reflect official positions of the Federal Reserve Bank of St. Louis, the Federal Reserve System, or the Board of Governors.

Federal Reserve Bank of St. Louis Working Papers are preliminary materials circulated to stimulate discussion and critical comment. References in publications to Federal Reserve Bank of St. Louis Working Papers (other than an acknowledgment that the writer has had access to unpublished material) should be cleared with the author or authors.

Photo courtesy of The Gateway Arch, St. Louis, MO. www.gatewayarch.com

TESTING ASSET PRICING MODELS WITH EULER EQUATIONS: IT'S WORSE THAN YOU THINK

September 1995

ABSTRACT

This paper reexamines the small sample properties of Hansen's (1982) Generalized Method of Moments (GMM) and Hansen and Jagannathan's (1989) estimation-free tests on simulated data from a more plausible consumption based asset pricing model. Previous studies are incomplete and misleading. A continuous distribution of consumption growth produces a near non-identification in the GMM criterion function, severe bias in coefficient estimates, misleading parameter confidence intervals even for very large samples and far worse overrejection problem in GMM tests of restriction than previously thought. Further, estimation-free methods advocated by Kocherlakota (1990) may also have very poor finite sample properties.

KEYWORDS: Generalized method of moments, Monte Carlo simulation, Markov Chain

JEL CLASSIFICATION: C1, E1

Christopher J. Neely
Economist
Federal Reserve Bank of St. Louis
411 Locust St.
St. Louis, MO 63102
(314) 444-8568-office
(314) 444-8731-fax
neely@stls.frb.org

1. INTRODUCTION

This paper reexamines the small sample properties of Hansen's (1982) Generalized Method of Moments (GMM) and estimation-free tests proposed by Hansen and Jagannathan (1989) on simulated data from a consumption based asset pricing model. In the environment considered here, both GMM and estimation-free tests have far poorer properties than previously imagined. In particular, these results qualify the recommendation by Kocherlakota (1990) who suggests estimation-free methods of judging asset pricing models may avoid the difficulties associated with GMM.

Much research in asset pricing has centered on the representative agent framework developed by Lucas (1978). Use of the Lucas framework permits the study of the relationship between movements in output and equilibrium asset prices in a one-good, pure exchange economy. There have been two main approaches to the use of representative agent models in the study of asset prices. The first approach is known as calibration and was popularized by Mehra and Prescott (1985). The second approach, exemplified by Hansen and Singleton (1982), is to estimate the model directly from the data and conduct formal econometric tests of overidentifying restrictions.

Calibration exercises aim to generate simulated data that have properties exhibited by data from the real world. Mehra and Prescott (1985), for instance, calibrated the model in an unsuccessful attempt to produce the high equity premium. More recently, Kandel and Stambaugh (1990) describe a model economy in which the distribution of

consumption growth is lognormal. The outcome of a Markov chain process determines the mean and variance of the distribution. The state of nature and consumption determine asset prices in this economy. Kandel and Stambaugh (1990) find the model generates simulated data which is consistent in important ways with data from the real economy.

The second approach is typified by Hansen and Singleton's (1982 and 1984) work describing the GMM for the estimation and testing of models using orthogonality conditions implied by stochastic Euler equations. They apply the GMM to the representative agent model using various monthly measures of consumption growth and asset returns data from the U.S. from 1959:2 to 1978:12. The data strongly reject the overidentifying restrictions on the model. As a limited information estimator, the GMM does not require a joint hypothesis about the nature of the underlying economy and the stochastic environment, and therefore is the most problematic of the rejections for the use of the representative agent framework.

The GMM rejections from formal tests of the model using real data would seem to make the data coming from simulated representative agent economies much less relevant. However, the size of the formal tests derived from estimators such as the GMM estimator or the estimation-free approach are based on asymptotic results and may be strongly prone to overrejection in small samples from these environments. Tauchen (1986) and Kocherlakota (1990) previously scrutinized the small sample properties of the GMM in the Mehra and Prescott (1985) asset pricing framework. Investigating the properties of two-stage GMM, Tauchen found that use of shorter lags in the instrument set produced nearly

asymptotically optimal parameter estimates and that the test of overidentifying restrictions performed well in small samples. Evidence that the small sample performance of the multi-stage estimators were superior to those of the two-stage estimators prompted Kocherlakota (1990) to investigate multi-stage GMM. Using different sets of parameter values, Kocherlakota (1990, p. 285) found that "assuming that the large sample properties of ... GMM estimators are true in small samples can lead one to 'overreject' the model."

GMM is only one member of the class of procedures that verify the restrictions implied by the stochastic Euler equations of the model. Another method is the estimation-free approach suggested by Hansen and Jagannathan (1989) in which the econometrician picks plausible values of the parameters a priori to see if they produce a "reasonable" fit.¹ Cochrane and Hansen (1992) used the estimation-free method (what they call "pricing error tests") as a point of comparison with the volatility tests of the equity premium puzzle. Comparing the properties of the estimation-free methods to GMM, Kocherlakota (1990) concluded (p. 287) "This paper...urges the use of Hansen and

¹ The Hansen and Jagannathan (1989) estimation-free methods studied here and in Kocherlakota (1990) should not to be confused with the variance bounds methodology explicated in Hansen and Jagannathan (1991). The finite sampling properties of the variance bounds methods have been investigated by Gregory and Smith (1992) and Burnside (1994).

Jagannathan's (1989) procedure, which tests the implication of each parameter specification separately."

There are two potential objectives for this paper. The first is to study the small sample properties of GMM and the estimation-free methods as econometric tests in a specific environment. The second is to study properties of the representative agent model, specifically whether the simulated data produce rejections in formal tests of the restrictions implied by the model. This paper focuses on the first objective.

This paper extends the previous studies of the GMM by Tauchen (1986) and Kocherlakota (1990) by examining in depth the small sample properties of both GMM and estimation-free tests in a more plausible asset pricing environment. The representative agent framework of Kandel and Stambaugh (1990) differs from those frameworks studied previously in that consumption growth has a *continuous* (rather than discrete) distribution and the calibrated model uses a much higher coefficient of relative risk aversion than previous works. The new, more plausible environment confirms some conclusions in the literature, but also makes it clear that they are incomplete and misleading in important ways. First, continuous consumption growth produces a near non-identification in the GMM criterion function that yields severely biased estimates and misleading confidence intervals for the parameters of interest even in very large samples. Also, the overrejection problem in GMM tests of model restrictions is far worse than previously thought, significantly overrejecting the model even with 8,000 observations for three of four estimators considered. The estimation-free tests of overidentifying

restrictions advocated by Kocherlakota (1990) may also have very poor finite sample properties, even worse than equivalent GMM estimators in some cases. Finally, the small sample properties of two-stage GMM are found to be better than those of multi-stage GMM in this environment.

2. THE SIMULATED ECONOMY

2.1 The Utility Function

The Kandel and Stambaugh (1990) model is based on that of Lucas (1978) who posited an infinitely lived, representative agent who maximizes expected time-additive utility, of the constant relative risk aversion class, subject to a budget constraint. The solution to this problem leads to the familiar Euler equation for each asset

$$\beta \cdot E_t \left[\left(\frac{p_{t+1} + d_{t+1}}{p_t} \right) \cdot \left(\frac{c_{t+1}}{c_t} \right)^{-\alpha} \right] - 1 = 0 \quad (1)$$

The parameter α is the coefficient of relative risk aversion; it measures the curvature of the utility function, the agent's tolerance for risk and the desire to intertemporally smooth consumption. The standard assumption of constant relative risk aversion ensures the equilibrium return process is stationary.

2.2 Consumption Growth

Aggregate consumption in each period is equal to aggregate output each period. The innovation in Kandel and Stambaugh's version of this model is that the distribution of consumption growth is continuous. The parameters depend on the state of nature which evolves according to a finite dimension, ergodic Markov process: $\ln(C_{t+1}/C_t)$ is normally distributed with mean μ_{t+1} and variance σ_{t+1}^2 , which are functions of the state of nature at time t .

2.3 The Asset Prices

The state of nature and consumption growth determine asset prices (hence asset returns) for each period through the Euler equation. Kandel and Stambaugh consider three types of assets: a riskfree bond, a share of aggregate wealth, and a share of levered equity. The payoff on the riskfree asset is one unit of the consumption good. The payoff to one share of aggregate wealth is a claim to all consumption in perpetuity. Levered equity is a share of aggregate wealth minus a claim on a risky bond. The Euler equation permits closed-form solutions for all three asset returns. The appendix provides more detailed information about the model economy, including the solutions to the asset pricing equations. They are also discussed more extensively in Mehra and Prescott (1985) and Kandel and Stambaugh (1988).

2.4 Calibrating the Model

Kandel and Stambaugh chose the parameters of their model economy in order to match the first two moments of consumption growth, the value-weighted New York Stock Exchange returns and the expected T-bill return from quarterly U.S. data. The parameter values were $\beta = .9973$, $\alpha = 55$, and $\theta = .478$.²

The value of $\alpha = 55$ seems extraordinarily high at first glance, but it is necessary to produce the desired equity premium and interest rate. Objections to such a high value of α are usually predicated on the results of thought experiments (Kandel and Stambaugh, 1988). For example, a value of $\alpha = 55$ means that a person with an income of \$50,000 would pay \$9,483 to avoid an even bet of \$10,000. The problem with these experiments is the assumption of *constant* relative risk aversion. Given appropriate sizes of a bet, almost any level of risk aversion could seem plausible or implausible. A value of $\alpha = 2$ is usually considered reasonable, but it means that a person with the same wealth would pay only \$1.25 to avoid an even bet of \$250. A person with $\alpha = 55$ would pay a more plausible \$33.96 to avoid the same bet. Kandel and Stambaugh argue the parameters should be chosen to match the data rather than ex ante expectations about the correct values of the parameters.

² The parameter θ governs the terms of the payoff to levered equity; it is discussed more fully in the appendix.

2.5 Results Using this Framework

Kandel and Stambaugh found that their simulated data exhibited the skewness and kurtosis typical of real consumption growth and asset return data. In addition, they were able to reproduce the "U" shaped pattern of autocorrelation of equity returns over return horizons using the equilibrium model of rational behavior described above. That is, the returns exhibited low negative first-order autocorrelation for returns at short horizons, more negative first-order autocorrelations at longer horizons and less negative first-order autocorrelations for returns at longer horizons.

Using a similar model of the economy, Cecchetti, Lam and Mark (1990, p. 398) point out that: "It is well known that serial correlation of returns does not in itself imply a violation of market efficiency. Nevertheless, there is a tendency to conclude that evidence of mean reversion in stock prices constitutes a rejection of equilibrium models of rational asset pricing." The rational asset pricing model constructed by Cecchetti, Lam and Mark (1990) produces data whose returns are negatively serially correlated. This illustrates that negative serial correlation in long-horizon stock returns "is consistent with an equilibrium model of asset pricing."³

³ Bonamo and Garcia (1994) argue that the Kandel and Stambaugh (1990) and Cecchetti, Lam and Mark (1990) results are due to an improperly calibrated model. Nevertheless, the basic environment is still of considerable interest.

3. THE GENERALIZED METHOD OF MOMENTS AND ESTIMATION-FREE METHODS

3.1 GMM Estimation

Intuitively, the idea behind GMM is to take a quadratic function of orthogonality conditions implied by a model and find parameter values which make sample counterparts of the orthogonality conditions close to zero, according to some optimal metric.⁴ In this case, the orthogonality conditions are those implied by the Euler equation:

$$E_{t-1} [f(R_t, \lambda_t, \pi)] = E_{t-1} [(\beta \cdot \lambda_t^{-\alpha} * R_t - 1) * Z_t] = 0 \quad (2)$$

where $\pi = \{ \alpha, \beta \}$ denotes the vector of parameters of interest, π^* the true values of those parameters, λ_t is consumption growth at time t , R_t is an N -vector of gross returns, Z_t is a K -vector of instruments and “*” denotes element by element multiplication. In practice the instruments used are generally lagged values of R , λ and a constant.

Define the function g_T to be the NK -vector of means of the sample orthogonality conditions

⁴ See Hansen (1982) for the original development of GMM and Hansen and Singleton (1982) for an application of GMM to estimate and test asset pricing models.

$$g_T(\pi) = \frac{1}{T} \cdot \sum_{t=1}^T f(R_t, \lambda_t, \pi) \quad (3)$$

Then the GMM chooses the parameters π to minimize a quadratic function of the sample orthogonality conditions

$$\hat{\pi} = \underset{\pi}{\operatorname{argmin}} g_T(\pi)' \cdot W \cdot g_T(\pi) \quad (4)$$

where W is an NK by NK weighting matrix of full rank. The optimal weighting matrix is the inverse of the variance-covariance matrix of the orthogonality conditions.

$$W = [E (f(R, \lambda, \pi) \cdot f(R, \lambda, \pi)')]^{-1} \quad (5)$$

Of course, (5) requires π to construct W , but W is used to construct the criterion function (4) that is minimized to find an estimate of π . In practice, we start by using an arbitrary matrix in (4), such as the identity matrix, to produce a consistent estimate $\hat{\pi}$ of π^* which can then be used in (5) to produce a consistent estimator of W^* , W_T . With the consistent estimator W_T , we can get an estimate of π^* that is both consistent and asymptotically efficient in its class. That estimator, known as the two-stage GMM estimator is

$$\hat{\pi}^* = \underset{\pi}{\operatorname{argmin}} g_T(\pi)' \cdot W_T(\hat{\pi}) \cdot g_T(\pi) \quad (6)$$

By repeatedly iterating over (4), (5) and (6) until the weighting matrix stops changing, we implement multi-stage GMM. Both two-stage and multi-stage GMM estimation are used in this paper.

3.2 Test of Overidentifying Restrictions

Hansen showed that under the null hypothesis that the model is true, the J-statistic given by

$$J_T = T \cdot [g_T(\hat{\pi}^*)' \cdot W_T(\hat{\pi}^*) \cdot g_T(\hat{\pi}^*)] \quad (7)$$

has an asymptotic chi-square distribution with $NK - 2$ degrees of freedom. The intuition behind the J-statistic is that if the model fits the data well, the sample counterparts of the orthogonality conditions can be made close to zero, the J-statistic will be small and we will be unable to reject the null hypothesis that the model is correct.

3.3 The Estimation-Free Test of Overidentifying Restrictions

If the model was correctly specified, then, under some regularity conditions on the pricing errors, the estimation-free statistic

$$J_T(\pi_0) = T \cdot [g_T(\pi_0)' \cdot W_0(\pi_0) \cdot g_T(\pi_0)] \quad (8)$$

would be asymptotically distributed as a chi-square random variable with NK degrees of freedom for the true parameter vector π_0 .

An estimation-free test of the model based on this statistic is performed by selecting a parameter specification, a priori, and comparing the resulting statistic (8) with the chi-square distribution with NK degrees of freedom. The distribution of this test statistic differs from that of the statistic produced by the GMM procedure because

parameter estimation pins down two orthogonality conditions to be degenerate random variables.

4. RESULTS USING DATA FROM THE REAL WORLD

Hansen and Singleton (1982) test the overidentifying restrictions implied by the representative agent asset pricing model using monthly consumption and returns data. For many of the combinations of instrument sets and asset returns they studied, they were able to reject the null hypothesis that the model was true. Because the GMM does not require a complete specification of the economy, this rejected the Kandel and Stambaugh model of the data generating process.

To facilitate the comparison to the Kandel and Stambaugh model economy which is calibrated for quarterly data, the null hypothesis was retested using two-stage and multi-stage GMM on quarterly data. The data were constructed from T-bill, consumption (nondurables and services) and population numbers taken from Federal Reserve data bases. Value-weighted New York stock exchange data were obtained from the Center for Research in Securities Prices (CRSP) tapes. Nominal returns were converted to real returns using the implicit consumption deflator. The data ran from 1959:1 to 1989:4, providing 120 quarterly observations.

Table 1 describes the six combinations of instrument sets and returns used on the quarterly data from the real world. The results from these combinations are shown in

Table 2. The initial values for the coefficients were those of the hypothesized model economy, $\beta = .99731$ and $\alpha = 55$.

Consistent with Tauchen (1986) and Kocherlakota (1990), the point estimates of α and β are quite sensitive to the choice of asset and/or instrument sets. The J-statistics frequently reject the model, particularly for those combinations that included T-bill returns. These results are consistent with those of Hansen and Singleton (1982), who applied GMM to monthly data.

5. GMM MONTE CARLO RESULTS FROM THE MODEL ECONOMY

5.1 Estimators

To determine the properties of GMM in the Kandel and Stambaugh environment, 1,600 samples of 90, 200, 500, 2,000, and 8,000 observations were drawn from the Kandel and Stambaugh representative agent model economy and the model was estimated by four GMM estimators on the simulated data. The estimators are distinguished by the asset returns and instrument set they use and by the number of iterations over the weighting matrix they permit. Table 3 describes the estimators used on simulated data. The first pair of estimators (TS1 and MS1) used only the riskless bond as an asset return, and a constant, lagged consumption growth and the lagged return on the riskless bond as instruments. The second set of two estimators (TS2 and MS2) used the riskfree bond, a share of aggregate wealth and a share of levered equity as the asset returns, and a

constant, lagged consumption growth and all lagged asset returns as the instruments.

TS1 and TS2 used two-stage GMM; MS1 and MS2 used multi-stage GMM.

A maximum of 75 iterations over the weighting matrix was permitted for the multi-stage GMM estimators. This constraint was not often binding for data sets of more than 90 observations. Also, a maximum of 200 iterations to numerically minimize the quadratic (equation (6)) was permitted. For all estimations, the starting values were the values of the parameters of the model economy, $\beta=.99731$ and $\alpha=.55$.

5.2 Small Sample Results

The results for a sample size of 90 observations are displayed in Table 4. The most prominent result confirms the finding of Kocherlakota that GMM is unreliable and there is a strong tendency for tests of restrictions from all estimators to "overreject" the model. Examining the last columns of Table 4, the actual rejection rates for the test of overidentifying restrictions are far higher in every case than the corresponding nominal size. In the case of estimator MS2, for example, the actual rejection rate is greater than 49 percent at a nominal 5 percent size.⁵ The overrejection is far stronger than that found by Kocherlakota (1990), whose maximum rejection rate for any estimator was 28 percent with 90 observations. Two reasons for the strong overrejections will be considered: poor

⁵ The maximum rejection rate of 61 percent was actually observed for the samples with 200 observations.

estimation of the parameters and the skewed and kurtotic distribution of the pricing errors.

Kocherlakota (1990) argued that the overrejection he found in his environment was due to poor estimation of the parameters. This proves to be an even greater problem in the Kandel and Stambaugh environment than it was in the Mehra and Prescott environment. The estimates for α are strongly biased downwards and but those of β are biased upwards; the confidence intervals for both parameters are highly misleading. The nominal 95 percent confidence intervals for α estimates cover the true value of α only 14.3 to 26.7 percent of the time. The coverage for the β confidence interval estimates are also poor, but not uniformly so, ranging from 31.1-92.8 percent. In contrast, the worst confidence interval performance found in Kocherlakota's environment was 41 percent for β and 47 percent for α . The results from this environment suggest GMM estimates of relative risk aversion will tend to grossly understate true risk aversion and confidence intervals will be far more misleading than previously thought.

Figure 1 illustrates the frequencies of the estimates of the pairs of parameters for the MS1 estimator for data sets of various lengths. The shape of the three dimensional histograms suggests a strong nonlinear relation between the parameter estimators. To investigate this relationship, numerical integration was used to construct the negative of the log of a simplified criterion function and its contour plot for the riskfree asset and the return to aggregate wealth. Figure 2 displays this function, that confirms the existence of

a very strong nonlinear (“U-shaped”) relation between the GMM parameter estimates in this environment.⁶

Investigation of the criterion function shows that the continuous distribution of consumption growth in the Kandel and Stambaugh environment is the culprit that generates the "U-shaped ridge." To see this, consider the simple criterion function in which we treat β and the return on the asset as constants for tractability.

$$CF_t = \beta^2 \cdot (E_{t-1}(CG_t^{-\alpha}))^2 \cdot R_t^2 - 2 \cdot \beta \cdot (E_{t-1}(CG_t^{-\alpha})) \cdot R_t + 1 \quad (9)$$

Recalling that consumption growth is lognormally distributed, using the moment generating function of a normal distribution and differentiating with respect to α , we get

$$\begin{aligned} \frac{\partial CF}{\partial \alpha} = & 2 \cdot \beta \cdot R_t \cdot (\beta \cdot R_t \cdot \exp(-\alpha \cdot \mu + .5 \cdot (-\alpha)^2 \cdot \sigma^2) - 1) \\ & \cdot \exp(-\alpha \cdot \mu + .5 \cdot (-\alpha)^2 \cdot \sigma^2) \cdot (-\mu + \alpha \cdot \sigma^2) \end{aligned} \quad (10)$$

In general, the first term in the expression will have more than one value of α that sets it equal to zero and those values will depend on the value of β . For instance, assuming $\mu = .0049$, $\sigma = .0128$, $\beta = .99731$ and $R_t = 1.025$, both $\alpha \approx 4.74$ and $\alpha = 55$ will set the first term on the right hand side of equation (10) to zero. The third term provides another

⁶ Figure 2 was constructed by numerical integration of the criterion function implied by the pricing errors of the riskless return and the return to aggregate wealth. The identity matrix was used as the weighting matrix. Alternatively, similar figures could be constructed using very large samples of simulated data with any of the estimators.

turning point; it is set to zero if $\alpha = \mu/\sigma^2 \approx 29$. If the variance of consumption growth were zero - that is, if it were discrete valued as in the Mehra and Prescott environment studied by Tauchen (1986) and Kocherlakota (1990) - the expectation of consumption growth in equation (9) would be a constant and the derivative with respect to α would be

$$\frac{\partial CF}{\partial \alpha} = 2 \cdot \beta \cdot R_t (\beta \cdot R_t \cdot CG^{-\alpha} - 1) \cdot (-\ln CG) \cdot CG^{-\alpha} \quad (11)$$

For a given value of β , there is only one value of α which sets (11) equal to zero and hence there is no "U-shaped ridge" in the Mehra and Prescott criterion function. There is, however, a more conventional straight ridge in this case.

Although the parameter estimates are biased and subject to near non-identification, it is not clear to what extent the parameter estimation is causing the overrejection. The use of estimation-free methods will shed more light on this issue.

The second factor in the overrejections may be the distributions of the pricing errors (the orthogonality conditions implied by equation (3)) which are highly skewed and kurtotic. (See Kocherlakota (1993) for a discussion of kurtosis in tests of asset pricing models.) If the pricing errors are kurtotic, the central limit theorem on which the J-statistic implicitly relies will not be a good approximation in small samples, there will be too many outliers and the statistic will tend to overreject.

In the simulated data, skewness and kurtosis are prominent features. For example, the coefficients of skewness and kurtosis constructed by numerical integration of the pricing error associated with the riskfree asset are approximately -3 and 21 respectively.

In contrast, the coefficients of skewness and kurtosis for a normal distribution are 0 and 3 respectively. The distribution of the pricing error implied by the riskfree asset is shown in Figure 3. Skewness and excess kurtosis are also prominent features of pricing errors in real data and could be contributing to rejections in the real data.

5.3 Small to Large Sample Results

Table 5 shows that the true sizes of GMM tests of overidentifying restrictions, median point estimates and confidence intervals for α (the coefficient of risk aversion) slowly converge to their asymptotic properties as the sample size increases. Disturbingly, rejection rates and the point estimates of and confidence intervals for β remain poor even with very large samples. In fact, the coverage rates for β can actually become poorer as the sample size increases. Examination of the contour map of the criterion function in Figure 2 exposes the source of the problem, however. As the sample size increases and the median α estimates converge on their true value of 55, the β estimates actually move along the U-shaped ridge away from the true value of β . The properties of the test of restrictions also do not move monotonically toward their asymptotic properties.

Kocherlakota (1990) noted that for some of his estimators, there is a trade-off between parameter estimation and the performance of the J-statistics. This seems to be the case in the Kandel and Stambaugh environment too, as reductions in the rejection rates may be accompanied by poorer confidence interval properties for the discount factor.

5.4 Two-Stage Versus Multi-Stage GMM

For both estimators and all samples sizes, the two-stage procedure has better properties in terms of the tests of overidentifying restrictions and parameter estimates than the multi-stage procedure. This contradicts work referred to in Kocherlakota (1990).

5.5 First Versus Second Estimators

The two estimators used in the Monte Carlo experiments were chosen to mimic a cross section of the estimators that provided strong rejections in the real data. The first estimators (TS1 and MS1) with only the riskless return and the smaller instrument set rejected less often in small samples, in contrast to real data. The second set of estimators (TS2 and MS2), that had all three asset returns and a larger instrument set, dominated the first set both with respect to correct size and coverage of the confidence intervals around parameter estimates as sample size increased. The two-stage procedure with the second estimator (TS2) has the best GMM performance in terms of rejection rates and estimates of the parameters.

Corroborating findings of Tauchen (1986) and Kocherlakota (1990), larger instrument sets make the estimation more imprecise; overrejection of the model is more likely for the second two estimators (TS2 and MS2) that use more information.

6. ESTIMATION-FREE MONTE CARLO RESULTS

To combat estimation problems influencing the test of overidentifying restrictions, Kocherlakota (1990) recommends the Hansen-Jagannathan (1989) estimation-free method to test restrictions implied by the model. Hansen and Jagannathan (1989) suggest treating each parameter specification as a different model and testing whether each specification of interest fits the data well. That is, a particular parameter set is chosen and the estimation-free J-statistic (8) is constructed from the orthogonality conditions implied by the data and the chosen parameter values. Because the parameters are chosen a priori, the test statistic is distributed as a chi-square random variable with NK degrees of freedom.

For each of the estimators, 1,600 random sample of various sizes were drawn and the estimation-free J-statistics were constructed with the true parameter values for various combinations of assets and instruments. Table 6 describes the estimators (combinations of orthogonality conditions) that were used in the investigation. The first estimator consisted of the risk-free return with a constant and lagged values of consumption growth and the risk-free return. The second estimator used all three asset returns with a constant and lagged values of consumption growth and the asset returns. The third estimator used the risk-free return with all instruments. The fourth and fifth estimator used combinations of asset returns with only a constant as an instrument.

Table 7 shows that the rejection rates from estimation-free tests of overidentifying restrictions greatly exceeded nominal sizes for all the estimators in moderately sized samples. Once again, larger instrument sets exacerbate the overrejection problem, but statistically significant levels of overrejection were observed for all the estimators. The overrejection problem in this environment appears to be much more severe than that examined by Kocherlakota (1990). In fact, the largest probability of rejection at a nominal 5 percent size observed by Kocherlakota was 10 percent, whereas estimator EF2 rejects 31 percent of the time at 5 percent size. This is still much less than an equivalent GMM estimator, MS2, which rejects 49 percent but it is more than the equivalent GMM estimator, TS2, which only rejects 24 percent of the time.⁷ Kocherlakota's prescriptive comment on GMM (p. 303), "The distribution of the J-statistic associated with this [estimation-free] procedure is more likely [than the GMM statistic] to be approximately chi-squared" is not generally true.

Figure 4 illustrates the overrejection problems and the convergence of the estimation-free test statistics to their asymptotic properties with the order statistics of the p-values from the observed J-statistics produced by estimator EF2, which uses all three asset returns and all five instruments. If the estimation-free J-statistics are truly chi-

⁷ A GMM estimator is said to be "equivalent" to an estimation-free method if they use the same combination of orthogonality conditions. Hence, estimation-free estimator EF1 is equivalent to TS1 and MS1 while EF2 is equivalent to TS2 and MS2.

square NK, the order statistics from their p-values should lie on the 45 degree line. P-values under the forty-five degree line indicate overrejection while the p-values that lie above the forty-five degree evidence underrejection. Although the asymptotic behavior is not as poor as that obtained using multi-stage GMM, there is a strong tendency to overreject in moderate sample sizes. Partly confirming the findings of Kocherlakota, estimation of the parameters does exacerbate the problem of severe overrejections in small samples.

7. CONCLUSIONS

Kandel and Stambaugh (1990) and Cecchetti, Lam and Mark (1990) have created models of the financial economy that illustrate some valuable lessons about the type of data consistent with rational asset pricing models. This paper has extended the literature on tests of asset pricing models by investigating the properties of GMM and estimation-free methods in such an environment. Confirming the results of Kocherlakota (1990), GMM procedures overreject the model and provide poor estimates of the parameters of interest. This work reveals that results from previous studies are incomplete and misleading in three ways. First, there is a near non-identification in the criterion function that induces very severe bias in estimates of the coefficient of relative risk aversion and misleading parameter confidence intervals even for very large samples. This near non-identification was not discovered in earlier studies because the authors

used a discrete rather than a continuous consumption growth specification. Second, the overrejection problem in GMM tests of model restrictions can be far worse than previously thought, more than 60 percent for one multi-stage estimator at a nominal 5 percent size. Finally, the estimation-free methods advocated by Kocherlakota (1990) may also have very poor finite sample properties, worse than those of two-stage GMM estimates for some estimators. The strong tendency to overreject the true model means that we should view the rejections of such models with real data with circumspection.

The near non-identification in the GMM criterion function produces a U-shaped ridge in the criterion function that leads to poor parameter estimation, including significantly biased estimates for α , consistent with Kocherlakota's results. Such bias supports Kandel and Stambaugh's claim that it is reasonable to contemplate a value of the coefficient of relative risk aversion that is substantially higher than commonly considered. In addition, because of the peculiar shape of the criterion function, the confidence intervals for the parameters remain misleading even at very large sample sizes.

Poor parameter estimation is not the whole story behind the overrejections of the model, as examination of the Hansen and Jagannathan estimation-free J-statistics show. Even with the true parameter values, the J-statistic still tends to overreject the model in small samples. Contrary to Kocherlakota's findings, however, some estimation-free J-statistics may be more prone to overrejection than their two-stage GMM counterparts. The overrejections of the model from the estimation-free methods must be caused by the

skewness and kurtosis of the pricing error data. This feature of the simulated data is shared by data from the real world.

APPENDIX A: THE KANDEL AND STAMBAUGH MODEL ECONOMY

A.1 The Utility Function

The agent chooses consumption and a portfolio of N assets in each period to solve:

$$\max_{c_\tau} E \left(\sum_{\tau=t}^{\infty} \beta^{\tau-t} \cdot \frac{c_\tau^{(1-\alpha)} - 1}{(1-\alpha)} \right), \quad (1)$$

where $0 < \beta < 1$ and $0 < \alpha < \infty$. The agent is subject to a budget constraint at time t :

$$c_t + p_t' q_t \leq p_t' q_{t-1} + w_t \quad (2)$$

where p_t , d_t and q_t are N -vector vectors of asset prices, dividends paid and asset quantities held at time t . w_t is labor income at time t . Maximization of (1) subject to (2) produces the following first-order necessary condition for each asset

$$1 = \beta \cdot E_t \left[\left(\frac{p_{t+1} + d_{t+1}}{p_t} \right) \cdot \left(\frac{c_{t+1}}{c_t} \right)^{-\alpha} \right] \quad (3)$$

A.2 Consumption Growth

Aggregate consumption in each period is equal to aggregate output each period. Denote consumption growth in period t by λ_t . The state of nature evolves according to a Markov process that determines the parameters of the continuous distribution of

consumption growth. That is, $\ln(\lambda_t)$ is normally distributed with mean μ_t and variance σ_t^2 , where μ_t and σ_t are a function of the state of nature at time t . There are s discrete states of nature which evolve according to a finite dimension, ergodic Markov chain. The transition matrix for the state of nature is denoted by ϕ (Φ).

Table A.1 - Markov Transition Probability Matrix for the State of Nature

State	1	2	3	4
1	.845	.113	.042	.000
2	.107	.750	.054	.089
3	.067	.067	.800	.067
4	.026	.079	.105	.789

Table A.2 - Parameters of the Distribution of Consumption Growth for Each State of Nature

State	μ	σ
1	.0049	.0128
2	.0041	.0128
3	.0049	.0131
4	.0041	.0131

In each period, the state of nature is chosen by a Markov process, then the realization of consumption growth for that period is drawn from the appropriate distribution (dependent on the realized value for the state). The state of nature and consumption growth determine asset prices (hence asset returns) for that period.

A.3 The Asset Prices

The asset prices in each period are given by functions of the state of nature and consumption growth. The infinite future discounted stream of asset dividends depend only on the current state and current consumption. Therefore, a sufficient statistic for the state of the economy in any period is the pair (c,i) , where c is the consumption in the period and i indexes the state of nature.

Kandel and Stambaugh consider three types of assets, a riskfree bond, a share of aggregate wealth, and a share of levered equity. Levered equity is a share of aggregate wealth minus a claim on a risky bond. To derive closed-form solutions for asset returns, one uses the property that the price of an asset is equal to the present value of the expected discounted future dividend stream and the Euler equation (3). The solutions to the asset pricing equations presented below are discussed extensively in Mehra and Prescott (1985) and Kandel and Stambaugh (1988). They will be presented below with minimal explanation.

The payoff on the riskfree asset is one unit of the consumption good. Substituting this in equation (3) for $p_{t+1} + d_{t+1}$ and using the property of the Markov transition matrix for the state of nature. It may be shown that the gross return on the riskfree asset when the economy goes from state i to state j is given by

$$R^f = \frac{1}{\beta \cdot \sum_{j=1}^s \phi_{ij} \cdot E[\lambda(i)^{-\alpha}]} = \frac{1}{\beta \cdot E[\lambda(i)^{-\alpha}]} \quad (4)$$

where ϕ_{ij} denotes the i th row, j th column of the Markov transition matrix for the states.

$\lambda(i)$ is consumption growth if the state of nature is i .

The payoff to one share of aggregate wealth is, of course, a claim to all consumption in perpetuity. Using the same sort of reasoning used to derive the riskfree return above, the return on aggregate wealth (going from state i to state j) is given by

$$R^{aw} = \lambda(i) \cdot \frac{(1 + w_j)}{w_i} \quad (5)$$

where w_i is the i th element of the s vector w given by

$$w = (I - H)^{-1} \cdot H \cdot 1_s \quad (6)$$

In equation (6), H is an $s \times s$ matrix with (i,j) th element

$$h_{ij} = \beta \cdot \phi_{ij} \cdot E[\lambda(i)^{1-\alpha}] \quad (7)$$

and 1_s is an $s \times 1$ vector of ones.

Levered equity is an asset consisting of a share of aggregate wealth net of a risky bond. That is, the holder of a share of levered equity has sold a bond to purchase a share

of aggregate wealth and will pay off according to the realization of the return to aggregate wealth. The payoff on a risky bond bought at time t is a fraction (θ) of aggregate wealth at time $t+1$ if aggregate wealth at time $t+1$ is greater than or equal to θ times the value of aggregate wealth at time t . If aggregate wealth at time $t+1$ is less than θ times the value of aggregate wealth at time t , then the risky bond holder receives all wealth at time t .

The return on levered equity is a complicated nonlinear function due to the nature of its risky payoff. Before defining the return on levered equity, first define the $s \times 1$ vector

$$g = Y \cdot 1_s \quad (8)$$

where the Y is the $s \times s$ matrix with (i,j)th element

$$y_{ij} = \beta \cdot \phi_{ij} \cdot E[\min[\lambda(i)^{1-\alpha} \cdot (1 + w_j), \lambda(i)^{-\alpha} \cdot \theta \cdot w_i]] \quad (9)$$

Then the gross return on levered equity may be given by

$$R^{le} = \frac{\max[0, \lambda(i) \cdot (1 + w_j) - \theta \cdot w_i]}{(w_i - g_i)} \quad (10)$$

REFERENCES

- Bonomo, M. and R. Garcia (1994), "Can a Well-Fitted Equilibrium Asset-Pricing Model Produce Mean Reversion?," *Journal of Applied Econometrics*, 9 (1), 19-29.
- Burnside, C. (1994), "Hansen-Jagannathan Bounds as Classical Tests of Asset Pricing Models," *Journal of Business and Economic Statistics*, 12 (1), 57-79.
- Cecchetti, S. G., P. Lam, and N. C. Mark (1990), "Mean Reversion in Equilibrium Asset Prices," *American Economic Review*, 80 (3), 398-418.
- Cochrane, J. H. and L. P. Hansen (1992), "Asset Pricing Explorations for Macroeconomics," *NBER Macroeconomics Annual*, 7, 115-165.
- Gregory, A. and G. W. Smith (1992), "Sampling Variability in Hansen-Jagannathan Bounds," *Economics Letters*, 38 (3), 263-267.
- Hansen, L. (1982), "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029-1053.

Hansen, L., and R. Jagannathan (1989), "Restrictions on Intertemporal Marginal Rates of Substitution Implied by Asset Returns," Working Paper 12, Northwestern University, Department of Finance.

Hansen, L., and R. Jagannathan (1991), "Implications of Security Market Data for Models of Dynamic Economies, " *Journal of Political Economy*, 99 (2), 225-262.

Hansen, L., and K. Singleton (1982), "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica*, 50 (5), 1269-1286.

_____ (1984), Errata, *Econometrica*, 52 (1), 267-268.

Kandel, S., R. F. Stambaugh (1988), "A Model of Expected Returns and Volatilities for Various Forecast Horizons," unpublished manuscript.

_____ (1990), "Expectations and Volatility of Consumption and Asset Returns," *The Review of Financial Studies*, 3 (2), 207-232.

Kocherlakota, N. R. (1990), "On Tests of Representative Consumer Asset Pricing Models," *Journal of Monetary Economics*, 26 (2), 285-304.

_____ (1993), "Testing the Consumption CAPM with Heavy Tailed Pricing Errors,"
University of Iowa Working Paper, 93-10.

Lucas, R. (1978), "Asset Prices in an Exchange Economy," *Econometrica*, 46 (6), 1429-1455.

Mehra, R., and E. Prescott (1985), "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 15 (2), 145-161.

Tauchen, G. (1986), "Statistical Properties of Generalized Method of Moment Models of Structural Parameters Obtained from Financial Market Data," *Journal of Business and Economic Statistics*, 4 (4), 397-425.

Table 1: Model Descriptions for GMM Estimation on Real Data

Estimator	Asset Returns	Instrument Set	Maximum Iterations Over W
TR1	T-bill	constant, CG_{-1} & $T\text{-bill}_{-1}$	2
MR1	T-bill	constant, CG_{-1} & $T\text{-bill}_{-1}$	75
TR2	VWNYSE	constant, CG_{-1} & $VWNYSE_{-1}$	2
MR2	VWNYSE	constant, CG_{-1} & $VWNYSE_{-1}$	75
TR3	T-bill & VWNYSE	constant, CG_{-1} , $T\text{-bill}_{-1}$ & $VWNYSE_{-1}$	2
MR3	T-bill & VWNYSE	constant, CG_{-1} , $T\text{-bill}_{-1}$ & $VWNYSE_{-1}$	75

Table 2: Results from GMM Estimation on Real Data

Estimator	sample size	d.f.	Alpha (s.e.)	Beta (s.e.)	J statistic (p-value)
TR1	120	1	-2.95 (0.71)	0.97 (0.00)	7.98 (0.00)
MR1	120	1	-2.94 (0.66)	0.97 (0.00)	8.70 (0.00)
TR2	120	1	-4.35 (5.49)	0.95 (0.03)	2.89 (0.09)
MR2	120	1	-4.58 (4.85)	0.95 (0.03)	2.33 (0.13)
TR3	120	6	1.23 (20.87)	0.97 (0.12)	5.99 (0.42)
MR3	120	6	-2.68 (0.57)	0.97 (0.00)	16.12 (0.01)

Estimator descriptions are provided in Table 1.

Table 3: Model Descriptions for GMM Estimation on Simulated Data

Estimator	Asset Returns	Instrument Set	Maximum Iterations Over W
TS1	Risk Free	constant, CG_{-1} & Risk Free ₋₁	2
MS1	Risk Free	constant, CG_{-1} & Risk Free ₋₁	75
TS2	Risk Free, Aggregate Wealth and Levered Equity	constant, CG_{-1} , Risk Free ₋₁ Aggregate Wealth ₋₁ & Levered Equity ₋₁	2
MS2	Risk Free, Aggregate Wealth and Levered Equity	constant, CG_{-1} , Risk Free ₋₁ Aggregate Wealth ₋₁ & Levered Equity ₋₁	75

Table 4: GMM Small Sample Results

Sample Size	Estimator	DF	Median Alpha	% of 95% confidence intervals covering true alpha	Median Beta	% of 95% confidence intervals covering true beta	Monte Carlo Critical Values			Model Rejection Rates		
							10%	5%	1%	10%	5%	1%
90	TS1	1	2.86	21.69	1.01	92.75	4.76	9.24	32.84	16.37	12.69	6.88
90	MS1	1	2.99	17	1.01	85.44	16.65	22.95	34.51	35.81	30.44	22.75
90	TS2	13	30.58	24.69	1.02	50.56	30.5	37.97	56.52	33.06	24.06	13.19
90	MS2	13	10.07	14.31	1.01	31.06	42.44	48.21	68.4	60.19	49.25	34.38

This table summarizes the results of applying GMM estimators, described in Table 3 to 1600 samples from the Kandel/Stambaugh model economy. The economy is described in Appendix A. Standard errors for the model rejection rates and confidence intervals are given by $\sqrt{(p(1-p))/1600}$.

Table 5: GMM Small to Large Sample Results

Sample Size	Estimator	DF	Median Alpha	% of 95% confidence intervals covering true alpha	Median Beta	% of 95% confidence intervals covering true beta	Model Rejection Rates		
							10%	5%	1%
200	TS1	1	5.45	27.13	1.02	87.38	19.63	14.56	9.19
500	TS1	1	10.89	37.75	1.03	70.81	20.50	15.06	9.75
2000	TS1	1	25.42	59.31	1.05	53.56	18.75	14.19	8.94
8000	TS1	1	40.78	78.81	1.04	61.75	13.44	8.31	3.75
200	MS1	1	5.52	18.81	1.02	77.25	39.75	34.44	27.69
500	MS1	1	10.91	29.37	1.03	56.88	37.75	31.94	25.19
2000	MS1	1	25.31	50.44	1.05	48.13	32.19	26.06	19.88
8000	MS1	1	40.84	73.94	1.05	59.00	22.50	18.00	10.94
200	TS2	13	37.01	31.81	1.03	54.56	33.44	23.56	11.69
500	TS2	13	40.98	48.56	1.04	59.38	20.31	13.44	5.44
2000	TS2	13	45.76	70.81	1.03	62.56	9.13	5.19	1.69
8000	TS2	13	51.58	86.06	1.01	81.44	8.25	4.06	0.69
200	MS2	13	7.89	15.19	1.01	29.44	68.06	60.75	48.50
500	MS2	13	14.89	20.38	1.02	26.44	60.00	53.75	43.50
2000	MS2	13	42.20	52.69	1.04	46.63	31.38	23.19	13.63
8000	MS2	13	51.20	83.06	1.01	77.50	14.88	8.88	2.69

This table summarizes the results of applying GMM estimators, described in Table 3 to 1600 samples from the Kandel/Stambaugh model economy. The economy is described in Appendix A. Standard errors for the model rejection rates and confidence intervals are given by $\sqrt{((p(1-p))/1600)}$.

Table 6: Model Descriptions for Estimation Free Methods on Simulated Data

Estimator	Asset Returns	Instrument Set
EF1	Risk Free	constant, CG_{-1} & Risk Free ₋₁
EF2	Risk Free, Aggregate Wealth & Levered Equity	constant, CG_{-1} , Risk Free ₋₁ , Aggregate Wealth ₋₁ & Levered Equity ₋₁
EF3	Risk Free	constant, CG_{-1} , Risk Free ₋₁ , Aggregate Wealth ₋₁ & Levered Equity ₋₁
EF4	Risk Free& Aggregate Wealth	constant
EF5	Risk Free, Aggregate Wealth & Levered Equity	constant

Table 7: Properties of the Estimation-Free Methods

Estimator	Sample Size	Model Rejection Rates at Nominal Size of		
		10%	5%	1%
EF1	90	16.25	10.13	3.41
EF1	200	13.38	7.41	2.63
EF1	500	10.63	5.78	1.41
EF1	2000	10.84	5.59	1.13
EF1	8000	9.97	5.28	1.06
EF2	90	43.38	31.44	15.75
EF2	200	43.44	32.88	15.81
EF2	500	30.94	21.19	9.88
EF2	2000	20.19	12.62	5.31
EF2	8000	13.06	7.44	1.94
EF3	90	14.42	8.56	2.33
EF3	200	17.46	10.75	3.08
EF3	500	13.06	7.42	2.02
EF3	2000	11.96	6.90	1.65
EF3	8000	10.94	5.10	1.38
EF4	90	14.58	9.44	3.48
EF4	200	12.04	6.77	1.92
EF4	500	10.23	4.98	1.33
EF4	2000	10.54	5.21	1.21
EF4	8000	9.54	4.79	0.96
EF5	90	16.25	10.13	3.41
EF5	200	13.38	7.41	2.63
EF5	500	10.63	5.78	1.41
EF5	2000	10.84	5.59	1.13
EF5	8000	9.97	5.28	1.06

This table summarizes the results of applying estimation free methods, described in Table 6 to 1600 samples from the Kandel/Stambaugh model economy. The economy is described in Appendix A. Standard errors for the model rejection rates and confidence intervals are given by $\sqrt{(p(1-p))/1600}$.

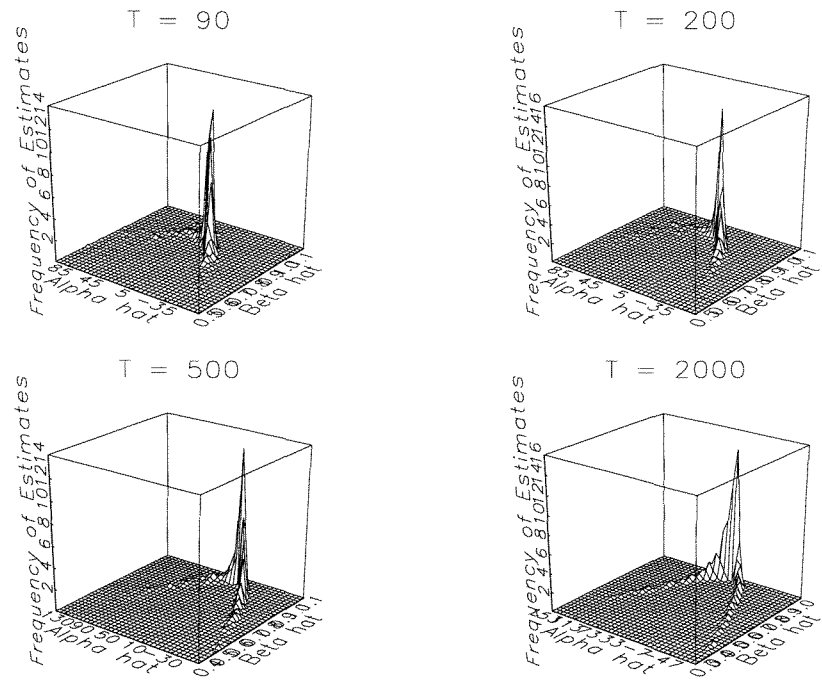


Figure 1: The frequencies of the parameter estimates from estimator MS1 on simulated data for various sample lengths. Note the strong nonlinear relationship between the parameter estimates.

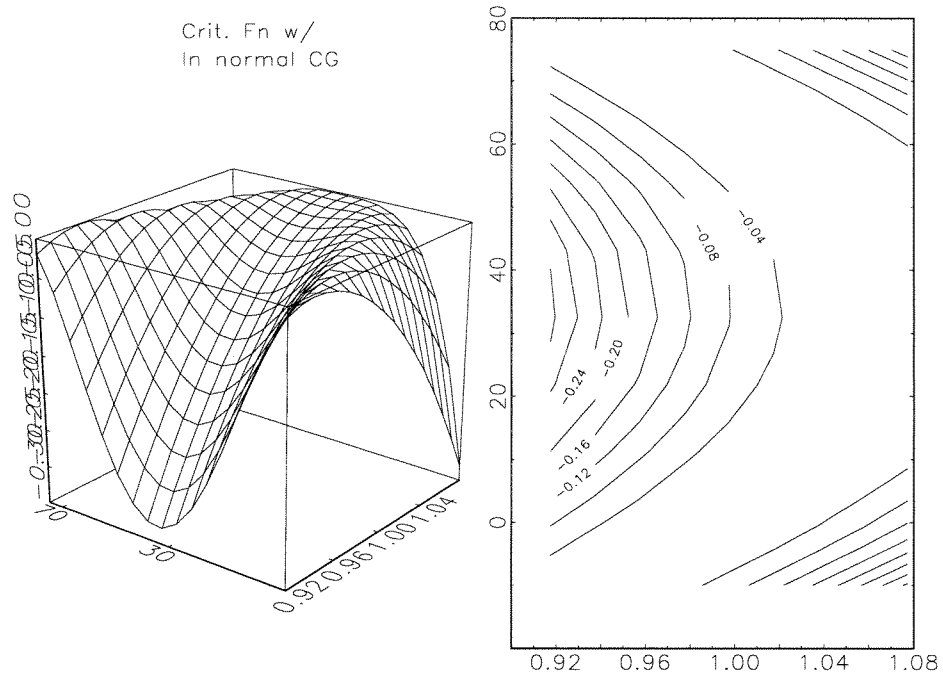


Figure 2: The log of the negative criterion function and its contour plot shows the "U-shaped" ridge creating the nonlinear relationship between the parameter estimates. The figure was constructed by numerical integration of the pricing errors implied by the risk free bond and the return to aggregate wealth.

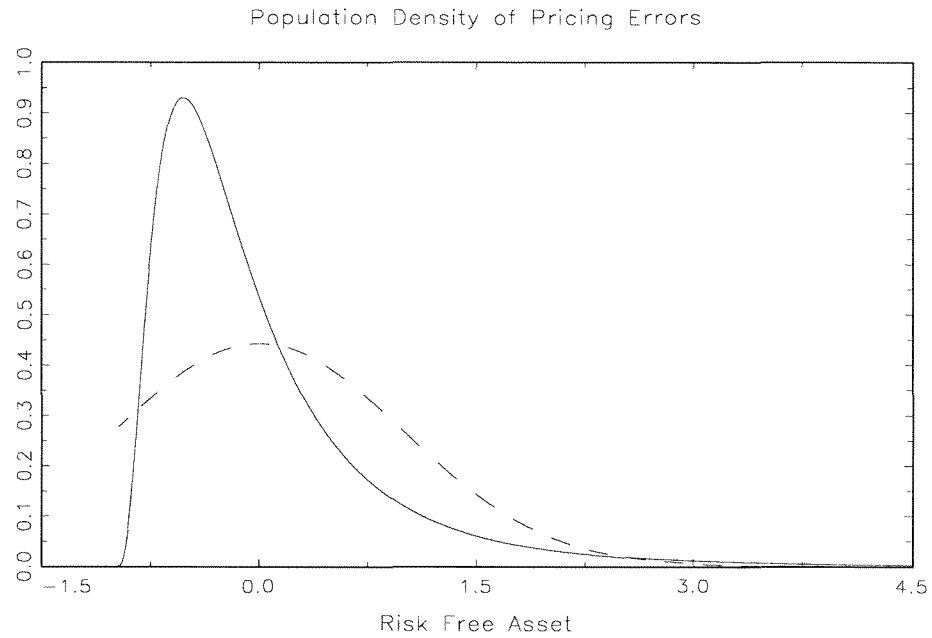


Figure 3: The population density of the pricing errors of the risk free asset overlaid by a normal density of the same mean and variance to highlight the skewness and kurtosis of the pricing error distribution. The distribution was constructed by numerical integration of the pricing error of the riskless bond.

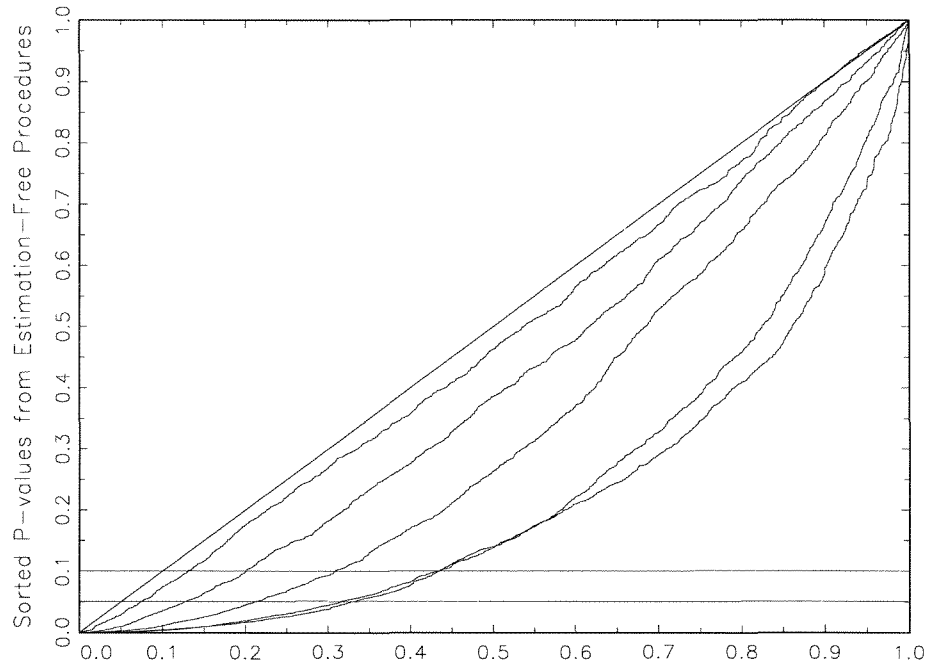


Figure 4: Sorted p-values from the "true" J-statistics, from estimator EF2, created from 1600 simulated data sets of length 90, 200, 500, 2000 and 8000 observations. The J-statistics are constructed from the "true" parameter values of the model economy. If the statistic is truly chi-square with NK degrees of freedom, the sorted p-values should lie along the 45 degree line. The horizontal lines at .05 and .1 may be used to find the degree of overrejection for each sample size. For example, the actual rejection rate for a sample size of 90 observations at a 10 percent nominal size is approximately 33 percent.